

BAB II

KAJIAN TEORI

A. Penilaian.

Penilaian merupakan hal sangat penting dalam proses pembelajaran.

Penilaian digunakan untuk mengambil keputusan penting terkait peserta didik, seperti menentukan apakah peserta didik tersebut perlu mengulang materi, naik kelas, mengulang atau tidak. Diperlukan pertimbangan yang matang untuk agar diperoleh keputusan yang tepat sehingga tidak merugikan peserta didik.

Untuk mendapatkan keputusan yang tepat, diperlukan informasi yang memadai tentang peserta didik, seperti penguasaan terhadap materi, sikap dan perilakunya. Dalam konteks ini penilaian memegang peranan yang cukup penting. Dari sini penilaian diharapkan memberi umpan balik yang objektif tentang apa yang telah dipelajari oleh peserta didik, bagaimana mereka belajar dan digunakan untuk mengetahui efektifitas dari proses pembelajaran.¹

Dengan demikian, apabila guru memiliki kemampuan untuk melakukan penilaian dengan baik maka dipastikan ia memiliki kemampuan mengajar yang baik pula. Uraian tersebut menandakan bahwa untuk menjadikan proses pembelajaran berkualitas maka guru seharusnya menguasai teknik penilaian yang baik pula. Sebab pembelajaran dan penilaian merupakan dua unsur yang tidak bisa dipisahkan dalam kegiatan belajar

¹ Kusaeri dan Suprananto, *Pengukuran dan Penilaian Pendidikan*, (Yogyakarta: Graha Ilmu, 2012) 8.

mengajar. Selanjutnya akan dijelaskan mengenai konsep dasar penilaian dan bagaimana cara memilih teknik penilaian yang tepat.

1. Konsep Dasar Penilaian.

Ada tiga istilah yang sering dipakai orang secara rancu, yaitu pengukuran, penilaian , dan evaluasi. Ketiga istilah ini memiliki arti yang sangat berbeda karena tingkat penggunaannya yang berbeda.²

Pengukuran merupakan cabang dari ilmu statitiska terapan yang bertujuan untuk membangun dasar-dasar pengembangan tes sehingga dapat menghasilkan tes yang valid dan reliabel. Arikunto mendefinisikan bahwa mengukur adalah membandingkan sesuatu dengan satu ukuran, pengukuran bersifat kuantitatif.³ Azwar mendefinisikan pengukuran sebagai suatu prosedur pemberian angka terhadap atribut atau variabel sepanjang kontinum.⁴ Dengan demikian, secara sederhana pengukuran dapat dikatakan sebagai suatu prosedur membandingkan antara atribut yang hendak diukur dengan alat ukurnya.

Penilaian lazimnya dimulai dari pengukuran. Menurut Gronlund & Linn penilaian adalah suatu proses yang sistematis dan mencakup kegiatan mengumpulkan, menganalisa, serta menginterpretasikan informasi untuk menentukan seberapa jauh seseorang mencapai tujuan pembelajaran yang ditetapkan.⁵ Jadi penilaian adalah suatu proses yang

² Kusaeri, *Acuan dan Teknik Penilaian Proses dan Hasil Belajar dalam Kurikulum 2013*, (Yogyakarta : Ar-Ruzz Media, 2014) 14.

³ Suharsimi Arikunto, *Evaluasi Program Pendidikan*, (Jakarta : Bumi Aksara, 2008) 3.

⁴ Azwar, *Sikap Manusia: Teori Dan Pengukurannya*, (Yogyakarta, Pustaka Pelajar, 1995) 3.

⁵ Gronlund & Linn, *Measurement And Evaluation In Teaching*, (New York, Mac Millan Publishing, 1990) 5.

sistematis dan mencakup kegiatan mengumpulkan, menganalisa, serta menginterpretasikan informasi yang dapat digunakan untuk membuat kesimpulan tentang karakteristik suatu objek berdasarkan baik dan buruk. Penilaian lebih bersifat kualitatif.

Evaluasi adalah kegiatan identifikasi untuk melihat apakah suatu program yang telah direncanakan telah tercapai atau belum, berharga atau tidak berharga, dan dapat pula untuk melihat tingkat efisiensi pelaksanaannya. Evaluasi adalah suatu proses penilaian untuk mengambil keputusan yang menggunakan seperangkat hasil pengukuran dan berpatokan kepada tujuan yang telah dirumuskan.⁶ Sehingga Kegiatan evaluasi dilakukan dengan cara membandingkan hasil pengukuran dan penilaian dengan suatu norma atau kriteria, dan hasilnya dinyatakan secara evaluatif.

Dalam penelitian ini definisi yang digunakan adalah penilaian. Karena yang diteliti adalah penilaian hasil belajar peserta didik. Untuk mengetahui tentang penilaian lebih lanjut maka akan dijelaskan bagaimana memilih teknik penilaian yang tepat.

2. Pemilihan Teknik Penilaian.

Ada beberapa alasan penting dalam pemilihan suatu teknik penilaian, agar hasil dari penilaian yang dilakukan benar-benar mendeskripsikan kemampuan dari peserta didik. Oleh karena itu berikut

⁶ Dadan Rosana, *Modul Evaluasi UT Bab I*, (Yogyakarta, 2011) 12.

disajikan bagaimana prinsip-prinsip dalam memilih teknik penilaian agar lebih bermakna.

Pertama, tujuan pembelajaran (dalam konteks sekarang dalam bentuk kompetensi dasar dan dirinci sebagai indikator). Sebelum menilai peserta didik guru harus mentukan tujuan pembelajaran. Semakin jelas dan spesifik tujuan pembelajaran semakin mudah dalam menentukan teknik penilaian yang tepat.

Kedua, teknik penilaian yang dipilih harus sesuai dengan tujuan pembelajaran. Misalnya, guru ingin menilai bagaimana peserta didik memecahkan masalah maka guru harus memilih teknik penilaian yang mampu untuk mengukur kemampuan pemecahan masalah yang dimiliki peserta didik.

Ketiga, teknik penilaian yang dipilih harus sesuai dengan karakteristik dan kebutuhan peserta didik. Pemilihan teknik penilaian yang tepat tidak hanya membantu peserta didik memperoleh informasi tentang hasil belajar namun juga akan sangat bermakna.

Keempat, dalam menginterpretasikan hasil penilaian guru harus mempertimbangkan kelemahan setiap teknik penilaian. Meskipun guru menggunakan teknik penilaian tertentu, informasi sebenarnya yang diperoleh adalah sebagian saja. Sehingga diperbolehkan menggunakan beberapa teknik penilaian untuk mengukur kemampuan peserta didik.⁷

⁷ Kusaeri, *Acuan dan Teknik Penilaian*, 24.

Dalam penitian ini, teknik penilaian yang digunakan adalah penilaian tes tertulis bentuk uraian non objektif. Jenis penilaian ini memiliki kriteria untuk mengukur kemampuan pemecahan masalah yang dimiliki peserta didik. Kriteria tersebut sesuai dengan pembelajaran berbasis masalah. Untuk lebih jelas akan dibahas mengenai tes tertulis bentuk uraian non objektif.

B. Tes Tertulis Bentuk Uraian Non Objektif.

Tes merupakan salah satu alat untuk melakukan pengukuran, yaitu alat untuk mengumpulkan informasi karakteristik suatu objek. Dalam pembelajaran, objek yang dimaksud adalah kecakapan peserta didik, minat, motivasi, dan sebagainya.⁸ Menurut Djemari Mardapi tes merupakan salah satu cara untuk menaksir besarnya kemampuan seseorang secara tidak langsung, yaitu dengan cara memberikan stimulus atau pertanyaan untuk mengetahui respon dari orang tersebut.⁹ Menurut Suharsimi Arikunto tes adalah suatu alat atau prosedur yang sistematis dan obyektif untuk memperoleh data-data atau keterangan yang diinginkan tentang seseorang, dengan cara yang boleh dikatakan tepat dan cepat.¹⁰ Menurut Kusaeri tes adalah suatu prosedur yang sistematis untuk menggambarkan karakteristik tertentu tentang peserta didik dengan menggunakan deskripsi dan angka.¹¹

Dari beberapa pendapat para ahli terkait pengertian tes, disimpulkan bahwa tes dalam kegiatan pembelajaran adalah serangkaian kegiatan yang

⁸ Eko Putro Widyoko, *Evaluasi Program Pembelajaran*, (Yogyakarta: Pustaka Pelajar, 2011) 45.

⁹ Djemari Mardapi, *Teknik Penyusunan Instrumen Tes Dan Non Tes*, (Yogyakarta : Mitra Cendekia, 2008) 67.

¹⁰ Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan*, (Jakarta : Bumi Aksara, 2013), 4.

¹¹ Kusaeri, *Acuan dan Teknik Penilaian*, 14.

dilakukan dengan cara memberikan pertanyaan-pertanyaan untuk mengetahui informasi-informasi terkait kemampuan peserta didik dalam memahami materi yang telah diberikan. Secara keseluruhan semua pendapat para ahli tentang pengertian tes memiliki kesamaan. Namun, secara lebih lanjut Kusaeri memberi penekanan bahwa hasil tes yang telah dilakukan berupa deskripsi dan angka.

Pada dasarnya untuk melakukan sebuah penilaian dapat digunakan dua bentuk instrumen, yaitu tes dan non tes. Instrumen tes meliputi tes tertulis bentuk pilihan dan uraian, sedangkan non tes terdiri dari portofolio, kinerja, proyek, penilaian diri, penilaian jurnal dan tes lisan.¹²

Menurut sejarah, tes yang pertama kali digunakan adalah tes tulis bentuk uraian. Tes tertulis bentuk uraian adalah Teori Tes Klasik atau *Classical True-Score Theory*, dinamakan Teori Tes Klasik karena unsur-unsur teori ini sudah dikembangkan dan diaplikasikan sejak lama, namun tetap bertahan hingga sekarang.¹³

Tes tertulis bentuk uraian merupakan seperangkat soal yang berupa tugas, pertanyaan yang menuntut peserta didik untuk mengorganisasikan dan menyatakan jawabannya menurut kata-kata sendiri. Jawaban tersebut dapat berbentuk mengingat kembali, menyusun, mengorganisasikan atau memadukan pengetahuan yang telah dipelajarinya dalam rangkaian kalimat

¹² Kusaeri, *Acuan Dan Teknik Penilaian*, 19.

¹³ Sumardi Suryabrata, "Pengembangan alat ukur psikologis," (Yogyakarta: Andi Offset, 2002) 21.

atau kata-kata yang tersusun secara baik.¹⁴ Sedangkan menurut Asmawi Zaenul dan Noehi Nasution, tes tertulis bentuk uraian adalah butir soal yang mengandung pertanyaan yang jawaban dari soal tersebut dilakukan dengan cara mengekspresikan pikiran peserta tes.¹⁵

Berdasarkan sistem penskorannya, tes tertulis bentuk uraian diklasifikasikan menjadi dua, yaitu tes tertulis bentuk uraian objektif dan non objektif. Tes objektif memberi pengertian bahwa penskorannya dilakukan secara objektif, karena bentuk soalnya menuntut sekumpulan jawaban dengan pengertian atau konsep tertentu. Sementara bentuk uraian non objektif menuntut jawaban berupa pengertian atau konsep berdasarkan pendapat masing-masing peserta tes, sehingga penskorannya sangat sulit untuk dilakukan secara objektif. Penskoran untuk tes tertulis bentuk uraian non objektif dinyatakan dalam bentuk rentangan.¹⁶

Eko Putro Widoyoko menambahkan bahwa penskoran tes uraian non objektif dipengaruhi oleh pemberi skor. Jawaban yang sama dapat memiliki skor yang berbeda oleh pemberi skor yang berbeda. Hal ini diakibatkan oleh beberapa hal, antara lain (a) Ketidak konsistenan penilai (b) *Haloo effect* atau kesan guru terhadap peserta didik sebelumnya (c) Pengaruh urutan pemeriksaan (d) Pengaruh bentuk tulisan dan bahasa.¹⁷

¹⁴ Kunandar, *Penilaian Autentik Penilaian Hasil Belajar Peserta Didik Berdasarkan Kurikulum 2013*, (Jakarta : Raja Grafindo Persada, 2014) 209.

¹⁵ Asmawi Zaenul, *Penilaian Hasil Belajar*, (Jakarta : Pusat Pengembangan Aktivitas Instruksional Ditjen Dikti, 2005) 37.

¹⁶ Kusaeri, *Acuan Dan Teknik Penilaian* 90.

¹⁷ Eko Putro Widyoko, *Evaluasi*, 47.

Namun untuk mengurangi efek dari faktor yang telah disebutkan oleh Eko Putro Widoyoko, hendaknya pedoman penskoran dibuat secara detail dan jelas. Hal ini dimaksudkan untuk mengurangi kesubyektifan penskoran dalam tes. Sehingga penskoran yang dilakukan untuk tes uraian non objektif menghasilkan data yang valid.

1. Kaidah Penulisan Tes Tertulis Bentuk Uraian Non Objektif.

Secara umum penulisan tes tertulis bentuk uraian non objektif harus memperhatikan beberapa hal sebagai berikut:¹⁸ (a) Mengukur kompetensi peserta didik. Artinya soal uraian tersebut mampu mengukur kemampuan peserta didik secara nyata dan akurat. (b) Soal uraian mampu mendorong peserta didik untuk berlogika dan berpikir tingkat tinggi. (c) Mengukur kemampuan berpikir kritis. (d) Materi yang diujikan hendaknya materi yang mampu merangsang kemampuan peserta didik untuk memecahkan masalah. (e) Pertanyaan yang diujikan hendaknya menggunakan kata kerja yang jelas dan mudah dipahami peserta didik. (f) Setiap soal harus mempunyai rubrik penskoran, dengan demikian hasil koreksi jawaban bisa lebih akurat.

Secara khusus penulisan tes tertulis bentuk uraian non objektif harus memperhatikan beberapa aspek berikut:¹⁹ Pertama, materi (1) Soal harus sesuai dengan indikator pada kisi-kisi. Artinya soal harus menyatakan perilaku dan materi yang hendak diukur sesuai dengan tuntutan indikator. (2) Batasan pertanyaan dan jawaban yang diharapkan

¹⁸ Kunandar, *Penilaian Autentik*, 211.

¹⁹ Kusaeri, *Acuan Dan Teknik Penilaian*, 92.

(ruang lingkup) harus jelas. (3) Isi materi sesuai dengan tujuan pengukuran. (4) Isi materi yang ditanyakan sesuai dengan jenjang, jenis sekolah dan tingkat kelas.

Kedua, konstruksi²⁰ (1) Rumusan kalimat soal atau pertanyaan harus menggunakan kata tanya atau perintah yang menuntut jawaban terurai; seperti : mengapa, uraikan, jelaskan, hubungkan, tafsirkan, buktikan, hitunglah, dsb. Jangan menggunakan kata Tanya yang tidak menuntut jawaban uraian, misalnya: siapa, dimana, kapan. Demikian juga jangan menggunakan kalimat tanya yang menuntut jawaban ya atau tidak. (2) Buatlah petunjuk yang jelas tentang cara mengerjakan soal. (3) Buatlah pedoman penyekoran segera setelah soal selesai ditulis dengan cara menguraikan komponen yang akan dinilai atau kriteria penskorannya, besarnya skor bagi setiap komponen, serta rentang skor yang dapat diperoleh untuk soal yang bersangkutan. (4) Hal-hal lain yang menyertai soal (grafik, tabel, gambar, peta, atau yang sejenisnya) harus jelas dan terbaca, sehingga tidak menimbulkan penafsiran yang berbeda.

Ketiga, bahasa²¹ (1) Rumusan kalimat soal harus komunikatif, yaitu menggunakan bahasa yang sederhana, dan menggunakan kata-kata yang sudah dikenal siswa, serta baik dari segi kaidah bahasa Indonesia. (2) Butir soal menggunakan bahasa Indonesia yang baik dan benar. (3) Rumusan soal tidak menggunakan kata/kalimat yang menimbulkan

²⁰ Kusaeri, *Acuan Dan Teknik Penilaian*, 92.

²¹ *Ibid.*, 93.

penafsiran yang berbeda (salah pengertian). (4) Rumusan soal tidak menggunakan kata-kata yang dapat menyinggung perasaan peserta didik.

2. Metode Pengoreksian Tes Tertulis Bentuk Uraian Non Objektif.

Untuk mengoreksi tes tertulis bentuk uraian non objektif dapat dilakukan dengan menggunakan metode *point method* dan *rating method*.²² *Point method* adalah metode pengoreksian dengan cara membandingkan setiap jawaban dengan jawaban ideal yang telah ditetapkan dalam rubrik penskoran. Skor yang diberikan kepada setiap jawaban akan tergantung pada derajat kepadanannya dengan rubrik penskoran.

Rating method adalah metode pengoreksian dengan cara setiap jawaban siswa ditetapkan dalam salah satu kelompok yang sudah dipilih-pilih berdasarkan mutunya selagi jawaban tersebut di baca. Kelompok-kelompok tersebut menyatakan mutu dan menentukan berapa skor yang dapat diberikan kepada setiap jawaban. Misalnya sebuah soal akan diberi skor maksimum 8, maka bagi soal tersebut dapat dibuat 9 kelompok jawaban dari 8 sampai 0.

Djemari Mardapi menambahkan bahwa untuk mengoreksi soal uraian hendaknya dilakukan dengan cara menilai jawaban pertanyaan demi pertanyaan bukan peserta didik ke peserta didik. Selanjutnya seorang guru menghilangkan identitas peserta didik dan menggantinya

²² Zainal Arifin, *Evaluasi Pembelajaran Prinsip, Teknik Prosedur*, (Bandung: Remaja Rosdakarya, 2013) 130.

dengan kode, hal ini bertujuan untuk menghindari adanya bias penilaian karena memiliki kesan baik atau jelek terhadap peserta didik.²³

Sedangkan menurut Kunandar, ada beberapa langkah untuk mengoreksi soal bentuk uraian non objektif agar mendekati objektif yaitu:²⁴ (a) menyusun pola jawaban yang diambil dari sampel jawaban peserta didik (b) pemeriksaan jawaban tidak dilakukan dengan cara mebaca jawaban satu peserta didik namun dengan cara pernomor (c) setiap lembar jawaban dikoreksi lebih dari satu kali (d) nilai peserta didik tidak langsung dijumlahkan secara global tetapi dirinci setiap aspek-aspek penilaian.

Dalam penelitian ini digunakan metode pengoreksian *point method*, dengan beberapa tambahan dari Djemari Mardapi. Hal ini bertujuan untuk meningkatkan keakuratan dari pengkoreksian yang akan dilakukan.

3. Penyusunan Pedoman Penskoran.

Pedoman penskoran merupakan panduan atau petunjuk yang menjelaskan tentang: Batasan atau kata-kata kunci untuk melakukan penyekoran terhadap soal-soal bentuk uraian dan kriteria-kriteria jawaban yang digunakan untuk melakukan penyekoran terhadap soal-soal bentuk uraian non-objektif. Dengan pedoman atau rubrik penskoran, guru dapat mengoreksi jawaban peserta didik secara akurat. Pedoman penskoran

²³ Djemari Mardapi, *Pengukuran, Penilaian Dan Evaluasi Pendidikan*, (Yogyakarta: Nuha Medika, 2012) 173.

²⁴ Kunandar, *Penilaian Autentik*, 246.

hendaknya disusun segera setelah perumusan kalimat butir-butir soal untuk menjaga keobjektivitasan dari penilaian yang akan dilakukan.²⁵

Rubrik penskoran diklasifikasikan kedalam dua bentuk, yaitu rubrik penskoran analitik dan holistik. (a) Rubrik penskoran analitik adalah rubrik penskoran dengan cara mengidentifikasi jawaban dari berbagai aspek yang berbeda. Skor untuk masing-masing aspek diletakkan secara terpisah.²⁶ (b) Rubrik penskoran holistik adalah rubrik penskoran dimana guru hanya memberikan skor tunggal berdasarkan pada keseluruhan jawaban peserta tes.

Dalam Penskoran analitik Djemari Mardapi menambahkan bahwa penskoran tersebut digunakan untuk soal ujian yang batas jawabannya sudah jelas dan terbatas. Misalnya soal mata pelajaran matematika dan fisika. Namun cara penskoran analitik juga bisa digunakan dalam bidang sosial dengan syarat batas jawabannya jelas dan komponen jawaban diberi skor.²⁷

Materi pelajaran fiqh merupakan materi yang jelas. Sehingga batas jawaban dalam pelajaran fiqh juga jelas. Untuk menjamin keakuratan penskoran terhadap tes yang dilakukan pada penelitian ini menggunakan pedoman penskoran analitik, karena pedoman penskoran analitik lebih detail bila dibandingkan dengan rubrik penskoran holistik.

²⁵ Kunandar, *Penilaian Autentik*, 244.

²⁶ Kusaeri, *Acuan Dan Teknik Penilaian*, 92.

²⁷ Djemari Mardapi, *Pengukuran, Penilaian* 173.

4. Keunggulan dan Kelemahan Tes Tertulis Bentuk Uraian Non Objektif.

Tes tertulis bentuk uraian non objektif memiliki beberapa keunggulan dan kelemahan. Keunggulan dari tes tertulis bentuk uraian non objektif adalah: ²⁸ (a) Mengukur aspek kognitif yang lebih tinggi. (b) Melatih kemampuan berpikir teratur pada peserta didik. (c) Mengembangkan ketrampilan memecahkan masalah. (d) Mengembangkan kemampuan berbahasa bagi peserta didik. (e) Penyusunan soal tidak membutuhkan waktu yang lama. (f) Menghindari sifat terkaan pada diri peserta didik. (g) Mampu memberikan gambaran yang tepat pada bagian-bagian yang belum dikuasai peserta didik.

Sedangkan kelemahan dari Tes tertulis bentuk uraian non objektif adalah sebagai berikut: (a) Sampel soal sangat terbatas sehingga bahan materi yang diujikan juga terbatas. (b) Cara memeriksa hasil tes sulit dan bisa mengandung unsur subyektivitas. (c) Membutuhkan waktu yang lama untuk proses koreksi. (d) Membutuhkan waktu yang lama untuk menyelesaikan satu soal uraian. (e) Tidak banyak kompetensi dasar yang dapat diuji.²⁹

²⁸ Kunandar, *Penilaian Autentik*, 213.

²⁹ *Ibid.* 214.

C. Tes Tertulis Bentuk Uraian Non Objektif Untuk Pembelajaran Berbasis Masalah.

Pembelajaran berbasis masalah atau disingkat dengan PBM adalah suatu model pembelajaran yang melibatkan peserta didik untuk memecahkan suatu masalah. Masalah yang digunakan adalah permasalahan yang ada pada dunia nyata, agar peserta didik mampu untuk belajar cara berpikir kritis dan keterampilan dalam pemecahan masalah, serta untuk memperoleh pengetahuan dan konsep yang esensial dari materi pelajaran.³⁰

Menurut Howard Barrows, masalah dalam pembelajaran berbasis masalah adalah masalah dalam dunia nyata yang disajikan secara mengambang (*ill-structured*). Pembelajaran berbasis masalah mampu untuk menunjang pembangunan kecakapan diri sendiri, kolaboratif dan kemampuan berpikir analisis, evaluasi dan mencipta.³¹

Untuk menunjang keberhasilan pembelajaran berbasis masalah, hendaknya menggunakan teknik penilaian yang tepat, agar kemampuan peserta didik dapat terukur. Tes tertulis bentuk uraian non objektif dapat digunakan untuk mengukur kemampuan hasil belajar peserta didik pada tingkat menganalisa, mengevaluasi dan mencipta. Atau dalam tingkatan

³⁰ Sudarman, *Problem Based Learning : Suatu Model Pembelajaran untuk Mengembangkan dan Meningkatkan Kemampuan Memecahkan Masalah*, (Jakarta, 2007, Dalam jurnal pendidikan inovatif).

³¹ Taufiq Amir, *Inovasi Pendidikan Melalui Problem Based Learning*, (Prenada Media Group : Jakarta, 2010) 22.

kemampuan berpikir C4, C5, C6.³² Karena dalam menjawab tes tertulis bentuk uraian non objektif peserta didik harus memulai dengan pengetahuan faktual yang dimilikinya dan mengorganisasikan fakta pilihannya dalam suatu susunan yang logis.

Kunandar juga menyatakan bahwa tes tertulis bentuk uraian non objektif dapat digunakan untuk mengukur kemampuan hasil belajar peserta didik pada tingkat C4, C5, C6. Karena tes tertulis bentuk uraian non objektif dapat menilai berbagai jenis kemampuan seperti: mengemukakan pendapat, berpikir kritis, berpikir kreatif dan pemecahan masalah.³³

Oleh karena itu, dalam merumuskan butir soal untuk tes tertulis bentuk uraian non objektif harus memperhatikan kemampuan peserta didik pada tingkat menganalisa, mengevaluasi dan mencipta.

1. Taksonomi Tujuan Pembelajaran.

Untuk melakukan penilaian yang baik maka perumusannya tidak bisa dipisahkan dari tujuan pembelajaran. Penilaian yang baik diturunkan dari tujuan pembelajaran yang jelas. Tujuan pembelajaran yang jelas akan sangat membantu agar penilaian yang dilakukan benar-benar mengukur apa yang telah diajarkan kepada peserta didik.³⁴

Tujuan pembelajaran yang baik memiliki indikator yang lengkap dan mencakup empat hal yaitu: *audience* (peserta didik), *behavior*

³² Masidjo, *Penilaian Pencapaian Hasil Belajar Siswa Disekolah*, (Kanisius: Yogyakarta, 1995) 46.

³³ Kunandar, *Penilaian Autentik*, 209

³⁴ Kusaeri, *Acuan Dan Teknik Penilaian*, 30.

(perilaku yang harus ditampilkan), *condition* (kondisi yang diberikan), dan *degree* (tingkatan yang diberikan).³⁵

Para ahli kurikulum telah sepakat untuk melakukan klasifikasi (taksonomi) tujuan pembelajaran. Terdapat bermacam-macam taksonomi tujuan pembelajaran, taksonomi tersebut diberi nama sesuai dengan nama penciptanya. Salah satu rujukan dalam sistem pendidikan nasional untuk merumuskan tujuan pembelajaran adalah Taksonomi Bloom.³⁶

Taksonomi pada dasarnya merupakan usaha pengelompokan yang disusun dan diurut berdasarkan ciri-ciri tertentu. Menurut Dadan Rosana taksonomi tujuan pembelajaran sangat diperlukan, karena pertimbangan sebagai berikut: (a) Perlu adanya kejelasan terminologi yang digunakan dalam tujuan pembelajaran, sebab tujuan pembelajaran berfungsi untuk memberikan arah kepada proses belajar dan menentukan prilaku yang dianggap sebagai bukti belajar. (b) Sebagai alat yang akan membantu guru dalam mendeskripsikan dan menyusun tes, teknik penilaian dan evaluasi.³⁷

Berdasarkan pertimbangan tersebut, diketahui bahwa taksonomi tujuan pembelajaran dapat membantu guru dalam penyusunan tes. Oleh karena itu, dalam proses penyusunan tes tertulis bentuk uraian non objektif untuk pembelajaran berbasis masalah yang baik, hendaknya

³⁵ Depdiknas, *Kurikulum Tingkat Satuan Pendidikan*, (Jakarta: Balitbang Depdiknas 2009) 14.

³⁶ Kusaeri, *Acuan Dan Teknik Penilaian*, 33.

³⁷ Dadan Rosana, *Modul Evaluasi UT Bab I*, (Yogyakarta, 2011) 46.

mengacu pada salah satu model taksonomi tujuan pembelajaran yang dikemukakan oleh para ahli.

Tes tertulis bentuk uraian non objektif merupakan tes yang bertujuan untuk mengukur kemampuan kognitif peserta didik. Maka dalam penelitian ini menggunakan Taksonomi Bloom edisi revisi domain kognitif. Taksonomi Bloom revisi dimensi proses kognitif yang berisikan enam kategori pokok, dengan jenjang yang paling rendah sampai jenjang yang paling tinggi. Selain domain kognitif dalam perumusannya juga memperhatikan dimensi pengetahuan yang meliputi pengetahuan faktual, konseptual, prosedural dan metakognitif.

Tingkatan berpikir Taksonomi Bloom edisi revisi adalah sebagai berikut: (a) Mengingat (*remember*) yaitu mengingat kembali pengetahuan yang relevan dari memori jangka panjang. (b) Memahami (*understand*) yaitu membangun pengetahuan dari pesan pembelajaran, termasuk komunikasi lisan, tertulis, dan grafis. (c) Menerapkan (*apply*) yaitu melaksanakan atau menggunakan prosedur dalam suatu situasi tertentu. (d) Menganalisis (*analyze*) yaitu memecah materi ke dalam bagian-bagian penyusunannya, dan menentukan bagaimana bagian-bagian tersebut saling berhubungan satu sama lain. (e) Mengevaluasi (*evaluate*) yaitu melakukan penilaian berdasarkan kriteria dan standar tertentu. (f) Menciptakan (*create*) yaitu menempatkan beberapa elemen secara bersamasama untuk membangun suatu keseluruhan yang logis dan

fungsi, dan mengatur elemen-elemen tersebut ke dalam pola atau struktur yang baru.³⁸

Kemampuan pada tingkatan menganalisis, mengevaluasi, dan mengkreasi merupakan kemampuan berpikir tingkat tinggi, sedangkan kemampuan pada tingkatan mengingat, memahami, dan menerapkan merupakan kemampuan tingkat rendah.³⁹ Pembelajaran berbasis masalah merupakan suatu model pembelajaran untuk melatihkan kemampuan berpikir tingkat tinggi peserta didik.⁴⁰ Oleh karena itu tes tertulis bentuk uraian non objektif yang disusun, hendaknya mengacu pada kemampuan berpikir tingkat tinggi pada level menganalisis, mengevaluasi, dan mengkreasi.

Kemampuan berpikir tingkat tinggi dibagi menjadi dua, yaitu berpikir kritis dan berpikir kreatif. Berpikir kritis merupakan kemampuan memberikan rasionalisasi terhadap sesuatu dan mampu memberikan penilaian terhadap sesuatu tersebut. Sedangkan berpikir kreatif adalah kemampuan untuk melakukan generalisasi dengan menggabungkan, mengubah atau mengulang kembali keberadaan ide-ide tersebut.⁴¹

Kemampuan peserta didik dalam berpikir kritis dan kreatif perlu

³⁸ Lorin Anderson and Krathwohl, *A taxonomy for Learning, Teaching, and Assessing*, (New York : Addison Wesley Longman, Inc, 2001) 67.

³⁹ Rini Julistiawati, *Keterampilan Berpikir Level C4, C5, & C6 Revisi Taksonomi Bloom Siswa Kelas X-3 Sman 1 Sumenep Pada Penerapan Model Pembelajaran Inkuiri*, *Journal of Chemical Education* Vol. 2 No.2 (Mei, 2013), 58.

⁴⁰ Wina Sanjaya, *Strategi Pembelajaran (Berorientasi Standar Proses Pendidikan)*, (Jakarta: Kencana, 2007) 218.

⁴¹ Kunandar, *Penilaian Autentik*, 171.

dilatihkan dan dikondisikan dengan baik melalui pembelajaran dan penilaian.

Dalam taksonomi bloom revisi juga diuraikan tentang klasifikasi dimensi pengetahuan dalam empat kategori, yaitu pengetahuan faktual, konseptual, prosedural dan metakognitif.⁴² Pengetahuan faktual berisikan pengetahuan tentang elemen dasar yang harus diketahui siswa untuk mengenal satu disiplin ilmu atau untuk menyelesaikan masalah didalamnya. Pengetahuan ini meliputi Pengetahuan tentang istilah dan pengetahuan tentang rincian dan unsur tertentu.

Pengetahuan Konseptual yaitu pengetahuan tentang hubungan timbal balik antara elemen-elemen dasar dalam suatu struktur yang memungkinkan elemen-elemen tersebut berfungsi secara bersama-sama. Pengetahuan ini mencakup Pengetahuan tentang klasifikasi dan kategori/ penggolongan, Pengetahuan tentang prinsip dan generalisasi, dan Pengetahuan tentang teori, model dan struktur.

Pengetahuan Prosedural yaitu pengetahuan tentang bagaimana melakukan suatu hal, metode dan inquiry, dan kriteria untuk menggunakan suatu keterampilan, algoritma, teknik dan suatu metode. Pengetahuan ini mencakup Pengetahuan tentang keterampilan dan algoritma tertentu, Pengetahuan tentang teknik dan metode tertentu dan Pengetahuan tentang kriteria untuk menentukan kapan menggunakan prosedur yang tepat.

⁴² Lorin Anderson, *A taxonomy*, 41-42.

Pengetahuan Metakognitif yaitu pengetahuan kognisi secara umum serta kesadaran dan pengetahuan tentang pengetahuan itu sendiri. Pengetahuan ini mencakup pengetahuan strategis, pengetahuan tentang tugas kognitif, termasuk pengetahuan kontekstual dan kondisional yang cocok, dan pengetahuan tentang diri sendiri.

2. Penyusunan Tes Tertulis Bentuk Uraian Non Objektif Untuk Pembelajaran Berbasis Masalah.

Dalam penyusunan tes tertulis bentuk uraian non objektif hendaknya memperhatikan beberapa hal penting untuk menjaga kualitas dari soal yang dikembangkan. Menurut Kunandar dalam penyusunan tes tertulis bentuk uraian harus memperhatikan hal-hal berikut: (a) pertanyaan hendaknya disusun untuk mengukur hasil belajar peserta didik yang tidak mungkin diukur dengan tes tertulis bentuk pilihan (b) pertanyaan hendaknya menuntut jawaban yang bersifat baru (c) menggunakan-kata-kata deskriptif (d) pertanyaan menggunakan bahasa yang komunikatif dan mudah dipahami (e) sebelum diujikan soal harus ditelaah terlebih dahulu.

Untuk menjamin keakuratan soal tes tertulis bentuk uraian non objektif, maka soal harus memenuhi syarat-syarat sebagai berikut: (a) membatasi ruang lingkup dengan memilih materi atau bahan pelajaran yang esensial (b) menggunakan bahasa yang baik dan benar sehingga mudah difahami dengan baik oleh peserta didik (c) jangan mengulang pertanyaan pada materi yang sama (d) tuliskan rubrik penskoran sebelum

menulis soal (e) menuliskan skor untuk masing-masing soal (f) rumusan soal harus jelas dan tegas (g) rumusan soal tidak boleh menggunakan kata yang menimbulkan penafsiran ganda (h) memiliki validitas yang tinggi (i) memiliki reliabilitas yang tinggi.⁴³

Untuk menghasilkan tes tertulis bentuk uraian non objektif yang berkualitas, dalam proses penyusunannya harus memperhatikan Taksonomi Bloom edisi revisi, aspek-aspek yang menjamin keakuratan suatu tes dan mengacu pada kaedah penulisan soal tes tertulis bentuk uraian non objektif yang telah disebutkan.

D. Analisis Kualitas Soal.

Alat ukur yang digunakan dalam penilaian hasil belajar harus dapat memberikan gambaran mengenai kemampuan belajar peserta didik yang sesungguhnya. Untuk itu, perlu dilakukannya analisis kualitas soal. Analisis soal adalah pengkajian pertanyaan-pertanyaan tes agar diperoleh seperangkat pertanyaan yang memiliki kualitas yang memadai.⁴⁴ Menurut Zainal Arifin Analisis kualitas tes merupakan suatu tahap yang harus ditempuh untuk mengetahui derajat kualitas tes, baik secara keseluruhan maupun butir soal yang menjadi bagian dari tes tersebut.⁴⁵ Dari beberapa pendapat tersebut disimpulkan bahwa analisis soal sangatlah penting guna mengetahui kualitas dari sebuah soal.

⁴³ Kunandar, *Penilaian Autentik*, 212.

⁴⁴ Nana Sudjana, *Penilaian Hasil Proses Belajar Mengajar*, (Bandung: Remaja Rosdakarya, 2013) 135.

⁴⁵ Zainal Arifin, *Evaluasi Pembelajaran*, 246.

Menurut Suharsimi Arikunto, suatu tes dikatakan baik sebagai alat pengukuran apabila memenuhi persyaratan tes. Persyaratan tes tersebut adalah validitas, reliabilitas, kepraktisan, obyektivitas, dan ekonomis.⁴⁶ Sedangkan menurut Wainer & Braun syarat penilaian yang bermutu adalah *valid, reliable and usable*.⁴⁷ Maka dalam penelitian ini akan dijelaskan mengenai validitas, reliabilitas dan kepraktisan:

1. Validitas.

Suatu tes dikatakan memiliki validitas yang tinggi apabila alat tersebut menjalankan fungsi ukurnya atau memberikan hasil ukur yang tepat sesuai dengan apa yang hendak diukur. Sumarna Supranata berpendapat bahwa “Validitas merupakan suatu konsep yang berkaitan dengan sejauh mana suatu tes dapat mengukur apa yang hendak diukur”. Validitas tes, secara keseluruhan ada empat macam validitas, yaitu: validitas isi (*content validity*), validitas konstruk (*construct validity*), validitas prediktif (*predictive validity*), dan validitas bandingan (*concurrent validity*).⁴⁸

Validitas isi sering dinamakan validitas kurikulum atau validitas kurikuler yang mengandung arti bahwa suatu tes dipandang valid apabila sesuai dengan materi yang ada dalam kurikulum. Untuk mengetahui apakah tes itu valid atau tidak, bisa dilakukan melalui penelaah kisi-kisi. Penelaah membandingkan kisi-kisi keseluruhan butir soal yang dibuat

⁴⁶ Suharsimi Arikunto, *Dasar-Dasar*, 8-11..

⁴⁷ Wainer & Braun, *Test Validity*, Hilldale: Lawrence Earlbaum Asociates, 1998) 20.

⁴⁸ Sumarna Supranata, *Analisis, Validitas, Reliabilitas dan interpretasi Hasil Tes.* (Bandung: Remaja Rosdakarya, 2005) 50.

dengan materi yang ada dalam kurikulum. Apabila sudah sesuai dipastikan soal tes tersebut mempunyai validitas isi yang baik.⁴⁹

Validitas konstruk menunjuk sejauh mana tes dapat mengukur dengan tepat aspek berpikir yang telah ditentukan dalam tujuan instruksional secara khusus.⁵⁰ Validitas konstruk dapat dilakukan dengan cara mencocokkan aspek-aspek berpikir dalam tes dengan aspek berpikir yang dikehendaki dalam tujuan intruksional khusus. Dalam hal ini, pengeraannya didasarkan pada logika. Selain itu, dapat juga dilakukan dengan cara melakukan diskusi dengan orang yang ahli di bidang yang bersangkutan.⁵¹

Validitas prediktif menunjuk pada kemampuan tes dalam meramalkan apa yang akan terjadi di masa mendatang. Dalam hal ini, kaitannya dengan prestasi hasil belajar peserta didik. Validitas prediktif dapat diketahui dengan mencari korelasi antar tes hasil belajar yang sedang diuji dengan kriteria validitas ramalan yang sudah ada. Jika kedua variabel menunjukkan korelasi yang signifikan, maka tes tersebut memiliki daya ramal yang tepat dalam artian pernah terjadi secara nyata dalam praktiknya.⁵²

Validitas bandingan menunjuk pada berapa jauh tes dapat mengukur tingkat penguasaan materi yang memang seharusnya dikuasai. Tes dikatakan memiliki validitas bandingan apabila tes tersebut dalam

⁴⁹ Djaali dan Puji Muljono, *Pengukuran dalam Bidang Pendidikan*. Jakarta: Grafindo, 2008) 50.

⁵⁰ Sudaryono, *Dasar-dasar Evaluasi Pembelajaran*, (Yogyakarta: Graha Ilmu, 2012) 142.

⁵¹ Suharsimi Arikunto, *Dasar-Dasar*, 83.

⁵² Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Raja Granfindo Persada, 2011) 170.

waktu yang sama menunjukkan hubungan searah antara tes pertama dengan tes berikutnya.⁵³

Validitas yang digunakan dalam penlitian ini adalah validitas logis yang meliputi validitas isi dan validitas konstruk. Validitas tersebut diperoleh dengan cara penilaian para ahli melalui proses validasi.

2. Reliabilitas.

Reliabilitas disebut juga tingkat atau derajat konsistensi suatu tes. Tes akan dikatakan reliabel apabila diperoleh hasil yang sama ketika suatu instrumen diteskan pada kelompok yang sama di waktu yang berbeda.⁵⁴

Tinggi rendahnya koefisien reliabilitas sangat dipengaruhi oleh beberapa faktor. Crocker dan Algina menyebutkan bahwa faktor itu antara lain adalah panjang suatu tes, kecepatan, homogenitas belahan, dan tingkat kesukaran soal. Tingkat kesukaran soal memegang peranan yang paling dominan.

Untuk mengetahui reliabilitas suatu tes bisa menggunakan mekanisme: teknik *test-retest*, belah dua, dan bentuk ekuivalen.⁵⁵ Sedangkan menurut Sumarna Surapranata ada empat konsep reliabilitas yaitu: paralel atau ekuivalen, *test retest*, belah dua, dan *internal consistency*. Namun sebagian ahli berpendapat bahwa metode belah dua

⁵³ *Ibid.*, 177.

⁵⁴ Zainal Arifin, *Evaluasi Pembelajaran*, 258.

⁵⁵ Djaali dan Puji Muljono, *Pengukuran*, 57.

merupakan bagian dari metode *internal consistency* sehingga pembagian metode menjadi tiga.⁵⁶

Teknik *test-retest* pengukurannya dilakukan dengan cara memberikan tes dua kali pada kelompok yang sama di waktu yang berbeda.

Bentuk tes ekuivalen merupakan dua buah tes yang dibuat setara seperti memiliki kesamaan tujuan, tingkat kesukaran dan susunan butir soal yang berbeda. Skor dari kedua kelompok tes dikorelasikan untuk mendapatkan reliabilitas soal.

Pengukuran reliabilitas teknik belah dua dengan mengorelasikan dua buah tes dari kelompok yang sama, membagi kedua tes tersebut menjadi dua bagian yang sama, kemudian mengorelasikan skor kedua belahan untuk mengestimasi reliabilitas tes. Menurut penelitian Aiken, tingkat kesukaran soal memegang peranan paling besar pada koefisien reliabilitas. Hal ini disebabkan karena menyangkut variasi jumlah soal yang dapat dijawab benar. Semakin sukar soal-soal dalam perangkat tes semakin besar pula variasi skor yang diperoleh belahan.

Penelitian ini menggunakan teknik belah dua. Kenyataannya, terdapat berbagai cara untuk membelah tes. Untuk perangkat tes dengan jumlah soal sebanyak enam, terdapat sepuluh cara dan sepuluh

⁵⁶ Sumarna Supranata, *Analisis*, 90.

kemungkinan estimasi reliabilitas seperti diungkapkan oleh Murphy dan Davidshofer dalam Tabel 2.1.⁵⁷

Tabel 2.1 Estimasi Reliabilitas Pembelahan Tes

Belahan Pertama	Belahan Kedua	Estimasi Reliabilitas
1,2,3	4,5,6	0,64
1,2,4	3,5,6	0,68
1,2,5	3,4,6	0,82
1,2,6	3,4,5	0,79
1,3,4	2,5,6	0,88
1,4,5	2,3,6	0,81
1,5,6	2,3,4	0,82
2,3,5	1,4,6	0,72
2,4,5	1,3,6	0,71
2,4,6	1,3,5	0,74

Reliabilitas yang digunakan pada penelitian ini adalah metode belah dua, dengan persamaan *Flanagan*. Belahan pertama terdiri dari soal nomor 1,4,5 dan belahan kedua terdiri dari soal nomor 2,3,6.

3. Tingkat Kesukaran Soal.

Tingkat kesukaran soal adalah peluang untuk menjawab benar suatu soal pada tingkat kemampuan tertentu, yang biasa dinyatakan dengan indeks. Indeks ini biasa dinyatakan dengan proporsi yang besarnya antara 0,00 sampai 1,00.⁵⁸

Menurut Djemari mardapi Mardapi, butir soal yang baik memiliki kisaran indeks kesulitan 0,3 – 0,7. Butir soal yang memiliki tingkat kesulitan di bawah 0,3 dianggap terlalu sulit dan butir soal yang memiliki

⁵⁷ Sumarna Supranata, *Analisis*, 104.

⁵⁸ Nana Sudjana, *Penilaian Hasil*, 225.

tingkat kesulitan di atas 0,7 dianggap terlalu mudah. Kriteria indeks daya beda butir soal yang boleh digunakan adalah $\geq 0,3$.⁵⁹

Menurut Crocker dan Algina tingkat kesukaran atau proporsi adalah nilai rata-rata dari kelompok peserta tes. Oleh karena itu tingkat kesukaran merupakan rata-rata dari suatu distribusi skor kelompok dari suatu soal.

Fungsi tingkat kesukaran butir soal biasanya dikaitkan dengan tujuan tes. Menurut Sukiman “Butir soal yang digunakan untuk keperluan ulangan atau ujian semester memiliki tingkat kesukaran yang sedang”. Indeks tingkat kesukaran butir soal yang baik antara 0,3- 0,7.⁶⁰

Jadi dalam penelitian ini instrumen penilaian hasil belajar yang dikembangkan harus memiliki indeks tingkat kesukaran soal pada rentang 0,3- 0,7.

4. Daya Pembeda Soal.

Yang dimaksud dengan daya pembeda soal adalah kemampuan suatu soal untuk membedakan antara siswa yang pandai dengan siswa yang kurang pandai. Logikanya adalah siswa yang pandai akan lebih mampu menjawab (mendapat skor lebih baik) dibanding dengan siswa yang kurang.⁶¹

Untuk menentukan daya pembeda pada soal uraian dilakukan dengan cara mengurutkan seluruh peserta tes berdasarkan perolehan skor total dari skor yang tinggi ke skor rendah. Setelah itu seluruh peserta tes

⁵⁹ Djemari Mardapi, *Teknik Penyusunan Instrumen Tes*, 143.

⁶⁰ Sukiman, *Pengembangan Sistem Evaluasi*. (Yogyakarta: Insan Madani, 2012) 201.

⁶¹ Daryanto, *Evaluasi Pendidikan*. 183.

dibagi menjadi 27% kelompok atas, yaitu kelompok yang memiliki skor total tinggi dan 27% kelompok bawah, yaitu kelompok yang memiliki skor rendah.⁶² Menurut Suharsimi Arikunto Butir soal yang baik adalah butir soal yang mempunyai indeks diskriminasi 0,3 sampai dengan 0,70.⁶³

Tindak lanjut butir soal sesudah dianalisis daya pembedanya sebagai berikut:⁶⁴

- a. Butir soal yang memiliki daya pembeda baik disimpan.
 - b. Butir soal dengan daya pembeda rendah, ada dua kemungkinan tidak lanjut yaitu: (1) ditelusuri untuk kemudian diperbaiki dan selanjutnya digunakan kembali dalam tes hasil belajar mendatang guna mengetahui daya pembedanya meningkat atau tidak. (2) Dibuang.
 - c. Butir item yang angka indeks diskriminasinya bertanda negatif, sebaiknya dibuang karena kualitas butir soalnya sangat jelek.

Dalam penelitian ini instrumen penilaian hasil belajar yang dikembangkan harus memiliki kriteria indeks daya beda soal pada rentangan 0,3 sampai 0,70.

⁶² Sumarna Surapranata, *Analisis, Validitas*, 31.

⁶³ Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan* (Jakarta: Bumi Aksara, 2013) 23.

⁶⁴ Anas Sudijono, *Pengantar Evaluasi*, 408.